Environmental Effects on Developing Wheat as Sensed by Near-Infrared Reflectance of Mature Grains

Stephen R. Delwiche,[1,2] Robert A. Graybosch,[3] Lenis A. Nelson,[4] and William R. Hruschka[1]

ABSTRACT

Cereal Chem. 79(6):885–891

For 30 years, near-infrared (NIR) spectroscopy has routinely been applied to the cereal grains for the purpose of rapidly measuring concentrations of constituents such as protein and moisture. The research described herein examined the ability of NIR reflectance spectroscopy on harvested wheat to determine weather-related, quality-determining properties that occurred during plant development. Twenty commercial cultivars or advanced breeding lines of hard red winter and hard white wheat (*Triticum aestivum* L.) were grown in 10 geographical locations under prevailing natural conditions of the U.S. Great Plains. Diffuse reflectance spectra (1,100–2,498 nm) of ground wheat from these samples were modeled by partial least squares one (PLS1) and multiple linear regression algorithms for the following properties: SDS sedimentation volume, amount of time during grain fill in which the temperature or relative humidity exceeded or was less than a threshold level (i.e., >30, >32, >35, <24°C; >80%, <40% rh). Rainfall values associated with four pre- and post-planting stages also were examined heuristically by PLS2 analysis. Partial correlation analysis was used to statistically remove the contribution of protein content from the quantitative NIR models. PLS1 models of 9–11 factors on scatter-corrected and (second order) derivatized spectra produced models whose dimensionless error (RPD, ratio of standard deviation of the property in a test set to the model standard error for that property) ranged from 2.0 to 3.3. Multiple linear regression models, involving the sum of four second-derivative terms with coefficients, produced models of slightly higher error compared with PLS models. For both modeling approaches, partial correlation analysis demonstrated that model success extends beyond an intercorrelation between property and protein content, a constituent that is well-modeled by NIR spectroscopy. With refinement, these types of NIR models may have the potential to provide grain handlers, millers, and bakers a tool for identifying the cultural environment under which the purchased grain was produced.

Despite considerable effort in research on the genetic basis of wheat (*Triticum aestivum* L.) quality (often defined by dough properties), the role of environment or seasonal variation has been much less studied and, consequently, is less understood. Whereas wheat quality is partially determined by genetics (hence the reliance on cultivar specification in many wheat-producing countries), cultural environment and its interaction with genotype can also determine quality (Busch et al 1969). In the United States, the environmental effect is often larger than the genetic effect on wheat quality (Peterson et al 1992). Such effects may include soil type, fertilizer level (especially N or S [Paredez-Lopez et al 1985; MacRitchie and Gupta 1993; Daniel and Triboi 2000; Luo et al 2000]), distribution of rainfall level (Faridi and Finlay 1989), and late season frosts (Lookhart and Finney 1984). In certain regions, elevated temperature during grain filling is possibly the most important environmental determinant of grain quality (Randall and Moss 1990). It is believed that high temperatures during grain filling, especially >35°C, alter the protein biosynthetic pathways of grain, leading to protein compositional changes (Blumenthal et al 1993). This temperature is particularly significant in light of wheat's general adaptation to moderate climates, such that the gradual increase in worldwide temperature from global warming may cause severe limitations for wheat cultivation (Ciaffi et al 1996). In a recent report by Blumenthal et al (1998), four hypotheses were proposed to account for changes in dough strength that are caused by heat stress: 1) changes in the ratio of glutenin to gliadin; 2) alteration in the formation of disulphide bonds between glutenin peptides, thus leading to a reduction of the size of the glutenin polymers; 3) the direct effects of heat-shock proteins on dough strength; and 4) changes that heat-shock proteins and chaperones impose on the folding and polymerization of polypeptides during polymer formation. In one of the very few studies on the effects of development temperature on the biochemical quality indices of North American wheats, Graybosch et al (1995) found that SDS sedimentation volume, a protein quality indicator that is relatively simple to measure, was an even more sensitive indicator of heat stress than either gliadin or glutenin content. Similar findings were reported for Australian wheat (Stone et al 1997). Although much simpler and faster to perform than HPLC analyses, the SDS sedimentation test, even if run in sets of parallel assays, can yield no more than 100 samples per laboratory day. A faster wheat quality test would allow breeders to screen more samples for genetic and environmental susceptibility to heat and other stress conditions. On a broad array of 30 North American wheat genotypes, we demonstrated that NIR reflectance equations could be developed for polymeric (considered to be glutenin) protein content, monomeric (gliadin) protein content, SDS sedimentation volume, and certain dough strength properties such as the mixograph time to peak, maximum resistance, and width of the mixing curve (Delwiche et al 1998). However, a problem that is endemic to protein-related NIR models is an inadvertent correlation between the modeled analyte and protein content itself, a constituent that is routinely measured by NIR. Various methods in NIR analysis have been used to uncouple the measurements of protein quality and protein quantity. Most notable are those reported by Wesley et al (1999, 2001) and Delwiche et al (1998). Whereas Wesley and coworkers reconstruct a spectrum through a curve-fitting procedure of summing component spectra (e.g., gliadin, glutenin) that are initially derived from a least squares curve fitting algorithm, the procedure of Delwiche is a statistical one in which partial correlation analysis is used on model residuals to remove the contribution of protein content. Because the relationship between heat stress and protein content is equivocal (Graybosch et al 1995; Stone and Nicolas 1995), successful NIR modeling will require the utilization of spectral absorbers that directly relate to the biochemical precursors of dough quality properties.

The objectives of the current research were to determine whether NIR reflectance is sensitive to biochemical properties that are influenced by the environment. Such properties are either directly measured, as in the case of SDS sedimentation volume, or represented by proxy, for example, by temperature- or humidity-dependent time periods of plant growth development. The intercorrelation of these properties with protein content is addressed in a manner similar to our previous research (Delwiche et al 1998).

MATERIALS AND METHODS

Wheat. Wheat was grown during the 2000 season as part of the Nebraska Winter Wheat Variety Tests program. Ten commercial cultivars or advanced breeding lines of hard red winter wheat (Alliance, Arapahoe, Cougar, Culver, Millennium, NE94654, NE95473, NE95510, Wesley, and 2137) and of 10 hard white (Betty, Heyne, Nuplains, NW97S154, NW97S182, NW97S218, NW97S278, NW97S312, NW97S343, and Trego) wheat were grown in field-replicated plots at each of 10 counties (Box Butte, Cheyenne, Dawes, Lincoln, Morrill, Perkins, Red Willow, Saunders, Scotts Bluff, and Webster) in Nebraska. Fertilizer (N, P, K) was applied according to standard practices at levels commensurate with soil fertilization needs. Though the tests program includes sites that possess irrigation, all sites chosen for the current study were dry land sites. At each location, or in close proximity to, state-administered weather stations recorded temperature, humidity, and precipitation values on an hourly basis throughout the growing the season. These stations are part of the Nebraska Automated Weather Data Network, maintained by the High Plains Regional Climate Center at the University of Nebraska, Lincoln. Samples were planted in late September through late October, with the exact planting date depending on the location. Likewise, samples were harvested in June through July of the following year. Dockage was removed from the samples before they were refrigerated ($\approx 0°C$) for 8–11 months until grinding.

Weather data was processed to form four temperature-related conditions, two humidity-related conditions, and four measures of rainfall for the pre- and post-planting stages of crop development. Temperature-related conditions pertained to the total number of hours during grain fill in which 1) the temperature exceeded 30°C [$t(T > 30°C)$], 2) the temperature exceeded 32°C [$t(T > 32°C)$], the temperature exceeded 35°C [$t(T > 35°C)$], and the temperature was less than 24°C [$t(T < 24°C)$]. Likewise, the total number of hours during grain fill in which the relative humidity was low [$t(<40\%\ rh)$] or high [$t(>80\%\ rh)$] was derived from the hourly weather data. Rainfall accumulation for the June-August preplanting period, the September-October planting period, the spring growing period up until anthesis, and the grain fill period also was determined.

Chemical Analysis. Samples from one field replicate at each of the 10 locations formed a set ($n = 198$) that was used in calibration equation development. Likewise, samples from the other field replicate constituted the test set ($n = 200$). Descriptive statistics of the biochemical and weather properties for the two sets are summarized in Table I. Field samples were split successively to produce a 20-g laboratory sample for grinding. A cylone grinder (Udy, Fort Collins, CO) equipped with a 0.5-mm screen was used to produce test samples for protein content, quality (SDS sedimentation), and NIR analyses. SDS sedimentation volume determination was performed according to Approved Method 56-70 (AACC 2000). Protein content ($N \times 5.7$) was determined using a combustion nitrogen analyzer (model FP-428, Leco Corp., St. Joseph, MI) on duplicate 150-mg portions of the test sample. Duplicate values were averaged. The error of the combustion
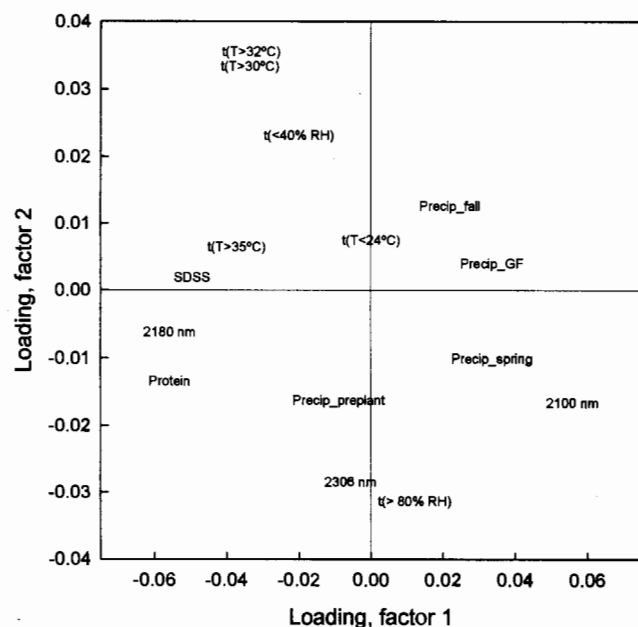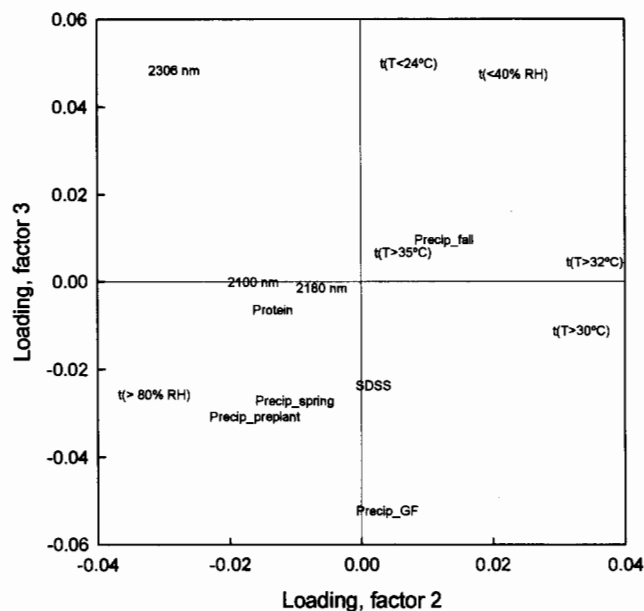

Fig. 1. PLS2 loadings for Factors 1 and 2, indicating relationship among protein content, SDS sedimentation volume, and environmental parameters plus four precipitation periods (preplant, fall, spring, and grain fill). Also included are the loadings for selected wavelengths ascribed to protein (2,180 nm), starch (2,100 nm), and oil (2,306 nm).

Fig. 2. PLS2 loadings for Factors 2 and 3 as described in Fig. 1.

TABLE I
Summary of Properties of Wheat Samples

Constituent or Property[a]	Calibration Set ($n = 198$)		Test Set ($n = 200$)	
	Range	Mean ± SD	Range	Mean ± SD
Protein content, %	9.09–18.27	13.35 ± 2.21	8.13–18.9	12.98 ± 2.26
SDSS, mL	10–35	21.2 ± 5.2	9–33	20.4 ± 5.6
$t(T > 30°C)$, hr	110–203	147.9 ± 30.4	110–203	148.0 ± 30.6
$t(T > 32°C)$, hr	33–114	72.8 ± 23.9	33–114	72.8 ± 24.1
$t(T > 35°C)$, hr	7–46	22.7 ± 10.9	7–46	22.8 ± 11.0
$t(T < 24°C)$, hr	792–1099	965.5 ± 89.3	792–1099	965.7 ± 88.9
$t(<40\%\ rh)$, hr	136–456	341.3 ± 113.4	136–456	341.0 ± 113.2
$t(>80\%\ rh)$, hr	165–518	364.2 ± 109.5	165–518	364.8 ± 109.1

[a] SDSS, SDS sedimentation volume; $t( )$, time during grain fill period when temperature or humidity was less than or greater than expression contained within the parentheses.

procedure, as defined by the standard deviation of single determinations of a check sample run in quadruplicate at the beginning and end of each of 11 analysis days throughout a one-month period, was 0.109% protein. Similarly, the error of the SDS sedimentation procedure, calculated from 30 measurements of a check sample during a one-month period, was 2.2 mL.

NIR Acquisition and Modeling. Test samples were conditioned for seven days in a constant humidity environment (33% rh) (Greenspan 1977) by placement in a desiccator that contained a saturated salt ($MgCl_2$) solution. The purpose of this conditioning was to minimize spectral differences between samples caused solely by moisture. NIR reflectance readings were collected at 2-nm increments over a wavelength range of 1,100–2,498 nm using a commercial scanning monochromator (model 6500 with spinning sample module, Foss-NIRSystems, Silver Spring, MD). Two packs of a standard forage cell (≈7 g material/pack) were scanned (32 scans/spectrum) and averaged. Samples from field reps 1 ($n = 198$) and 2 ($n = 200$) formed the calibration and test sets, respectively. To reduce spectral variation caused by sample-to-sample differences in particle size distribution or packing density, a multiplicative scatter correction transformation, as defined in Martens and Naes (1989), was applied to both sets, with the calibration set mean spectrum as the reference. To accentuate absorption peaks, a Savitzky-Golay second-derivative transformation (determined by fitting a quadratic polynomial to a wavelength and its three closest neighbors on each side, evaluating the polynomial at the central wavelength, then repeating this task for all wavelengths) was applied to each spectrum before partial least squares analysis.

Commercial software (Unscrambler v. 7.6, Corvallis, OR) was used to develop partial least squares (PLS1 and PLS2) multivariate models. As discussed in Martens and Naes (1989), PLS1 is a least squares regression procedure that initially reduces the dimension of the wave-length space from the number of wavelength points originally stored (the X block, 700 points in the present case) to a handful (typically 1–20) of orthogonal factors, while simultaneously incorporating the influence of a dependent (Y block) variable. PLS2 modeling differs from PLS1 in that more than one dependent variable is contained in the Y block. During model development, all Y variables may simultaneously influence the compression of the X block. PLS2 modeling is used in an explorative phase, whereby it becomes possible to graphically reveal the relationship among X and Y variables. For this study, the Y block consisted of 12 measurements: protein content, SDS sedimentation volume, the four temperature-related time values, the two humidity-related time values, and the four precipitation values. Before PLS2 analysis, each of the 700 X block variables and each Y block variable was scaled to equal variance by dividing by its standard deviation (i.e., scaling weight = 1/S), as determined from all calibration set samples.

PLS1 modeling (1,114–2,484 nm) was used to develop a calibration equation for each of the Y block variables, excluding the precipitation variables. One-sample-out cross validation was applied to determine the optimal number of factors, based on an *F*-test ($\alpha = 0.25$) using the minimum root mean squared differences (RMSD) and an RMSD from a model employing fewer factors. Each calibration model was subse-

TABLE II
Correlation Between Selected Environmental Properties and Protein Content or SDS Sedimentation Volume (SDSS)

Property[a]	Pearson Correlation Coefficient (r)[b]	
	Protein Content	SDSS
SDSS	0.903	· · ·
t(T > 30°C)	0.485	0.456
t(T > 32°C)	0.433	0.403
t(T > 35°C)	0.445	0.436
t(T < 24°C)	0.000 (ns)	−0.089 (ns)
t(<40% rh)	0.249	0.210 ($P = 0.003$)
t(>80% rh)	−0.286	−0.268

[a] t( ), time during grain fill period when temperature or humidity was less than or greater than expression contained within the parentheses.
[b] All correlations are significant at $P < 0.001$ unless otherwise noted; ns, not significant at $P = 0.05$.

Fig. 3. PLS2 scores for Factors 1 and 2. Symbol size is linearly related to protein content (smallest and largest sizes 9.09 and 18.27%, respectively).

TABLE III
Summary of PLS1 NIR Model Performances[a]

Constituent or Property[b]	Cross Validation of Calibration Set ($n = 198$)		Test Set Statistics ($n = 200$)[d]				
	PLS Factors	RMSD[c]	Bias	r	$r_{partial}$	SEP	RPD
Protein content, %	9	0.12	−0.01	0.999	· · ·	0.12	18.9
SDSS, mL	9	2.6	0.2	0.931	0.555	2.0	2.7
t(T > 30°C), hr	11	6.2	−2.9	0.952	0.950	9.4	3.3
t(T > 32°C), hr	11	5.4	−2.2	0.938	0.937	8.3	2.9
t(T > 35°C), hr	11	4.5	0.3	0.870	0.835	5.4	2.0
t(T < 24°C), hr	11	24.5	5.0	0.949	0.949	29.1	3.1
t(<40% rh), hr	12	35.4	13.9	0.917	0.912	45.4	2.5
t(>80% rh), hr	12	32.0	−16.7	0.935	0.929	38.7	2.8

[a] Model conditions: PLS1 on second-derivative (Savitzky-Golay, 7 point [14 nm] convolution window, second-order polynomial) of multiplicatively scatter-corrected (first-order regression to mean of calibration set spectra) spectra.
[b] SDSS, SDS sedimentation volume; t( ), time during grain fill period when temperature or humidity was less than or greater than expression contained within the parentheses.
[c] Root mean squared differences from a one-sample-out cross validation.
[d] Bias, mean of NIR model values minus mean of measured values; r, correlation coefficient of NIR-predicted and measured values; $r_{partial}$, partial correlation coefficient adjusted for protein content; SEP, standard error of performance (standard deviation of residuals); RPD, standard deviation of measured values divided by SEP.
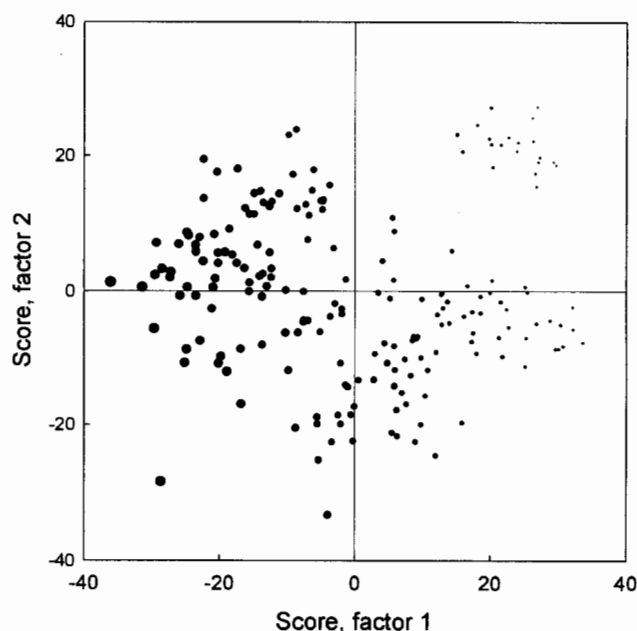
quently applied to the test set samples, whereupon model predictions were compared with actual measured values. Such comparisons were summarized by the standard error of performance (SEP), the correlation coefficient (r), the bias, and the ratio of the standard deviation of the measured values to the SEP, also known as the RPD (Williams and Sobering 1993). Additionally, the contribution of the spectral sensitivity to protein content in each of the other PLS1 models was statistically removed through the application of partial correlation analysis, as described in Fisher and van Belle (1993).

An in-house computer program was used to perform stepwise multiple linear regression. This program is designed to search for the single best second-derivative wavelength, starting with a derivative gap size (half-width of the convolution interval) of 10 nm. The gap size is then incremented by 2 nm, whereupon the search for the best wavelength resumes. Gap size incrementing and searching continues until a prescribed upper value (48 nm in the present research) of the gap is reached. The single best term is selected to be the combination of gap and wavelength that yields the highest coefficient of determination. Having selected the best single term, a second term is added and the cycling of second derivative gap size trials for the new term

resumes. Upon finding the best second term, the user may then fine-tune the selection of the first term by repeating the wavelength search and gap cycling, while holding the second term constant. Continued cycling between the holding of one term constant while searching for the other term is repeated until both terms have stabilized to yield a local optimum value in r^2. Third and higher terms are selected in a similar fashion, with each additional wavelength adding a significant order of complexity to the fine-tuning procedure. In the present study, up to four second-derivative terms were selected for each Y variable's regression equation.

RESULTS AND DISCUSSION

Loadings for the first three factors from the PLS2 analysis are shown Figs. 1 (2 vs. 1) and 2 (3 vs. 2). All Y block variables are plotted, in addition to specific wavelengths (X block), which from prior knowledge (Osborne and Fearn 1986) are known to be associated with absorption peaks for protein (2,180 nm), starch (2,100 nm), and oil (2,306 nm). Each point represents the contribution of that attribute (constituent or property in the case of a Y block variable,

TABLE IV
Summary of Multiple Linear Regression NIR Model Performances[a]

Constituent or Property[b]	Wavelengths (nm)	Gaps (nm)	Coefficients[c]	Bias	r	$r_{partial}$	SEP	RPD
				Test Set Statistics ($n = 200$)[d]				
Protein content, %	1506	34	−663.2	0.00	0.998	· · ·	0.13	17.1
	1606	38	541.5					
	1688	46	−351.7					
	2170	22	−806.9					
			17.35					
SDSS, mL	1660	22	−3122	0.1	0.940	0.616	1.9	2.9
	2182	22	−5250,					
	2246	10	−4127					
	2274	12	−11,970					
			−27.2					
t(T > 30°C), hr	1812	14	100200	−1.1	0.938	0.918	10.8	2.8
	1832	10	−380300					
	1832	12	584600					
	1870	10	−39880					
			576					
t(T > 32°C), hr	1270	12	−414800	−1.0	0.900	0.876	10.6	2.3
	1700	10	67900					
	1826	14	−59120					
	1832	14	156200					
			320					
t(T > 35°C), hr	1230	10	144900	−1.0	0.782	0.727	7.0	1.6
	1544	12	−67610					
	1704	24	20740					
	2226	12	45470					
			−209					
t(T < 24°C), hr	1206	10	−565400	9.6	0.850	0.851	46.8	1.9
	1288	10	−1703000					
	1544	12	509200					
	1830	16	378400					
			−190					
t(<40% rh), hr	1542	10	649000	−3.2	0.709	0.686	80.6	1.4
	1610	48	−146700					
	2154	30	−243800					
	2444	40	140000					
			89.7					
t(>80% rh), hr	1112	10	−1120000	0.5	0.797	0.777	66.0	1.6
	1270	12	−2736000					
	1544	12	−954400					
	1836	18	218300					
			986					

[a] Model conditions: four wavelengths of second-derivative of multiplicatively scatter-corrected (first-order regression to mean of calibration set spectra) spectra.

[b] SDSS, SDS sedimentation volume; t( ), time during grain fill period when temperature or humidity was less than or greater than expression contained within the parentheses.

[c] Constant term in the regression equation in bold type.

[d] Bias, mean of NIR model values minus mean of measured values; r, correlation coefficient of NIR-predicted and measured values; $r_{partial}$, partial correlation coefficient adjusted for protein content; SEP, standard error of performance (standard deviation of residuals); RPD, standard deviation of measured values divided by SEP.

transformed spectral response at a given wavelength in the case of an X block variable) to the first two PLS factors. Negative loading values for Factor 1 occurred for protein content, SDS sedimentation volume, and the three temperature-dependent time properties that quantify the time above the threshold temperatures of 30, 32, and 35°C. A second-derivative absorption peak at 2,180 nm (amide I and amide III combination band) possessed a negative Factor 1 loading value of approximately the same magnitude as those for protein and SDS sedimentation volume. Conversely, the loading value for the peak at 2,100 nm (starch O-H and C-O combination) was approximately equal in magnitude but of opposite sign to that at 2,180 nm, while the peak at 2,306 (oil CH_2 stretch-bend combination) possessed a very small Factor 1 loading value. This is supportive of the complementary behavior of starch and protein and a much smaller roll of lipids in the ground meal. Of the four precipitation parameters, three possessed Factor 1 loading values that were of opposite sign to the loading value for protein content, while the magnitude of the fourth parameter's (Precip_preplant) loading value was close to zero. Analysis of the PLS2 loadings revealed that more than half (53.5%) of the total

variation of the 12 Y block values were explained by four factors. At 10 factors, this value rose to 81.8%.

The degree of correlation between the temperature- or humidity-dependent time properties and protein content or SDS sedimentation volume is summarized in Table II. The greatest correlation occurred between SDS sedimentation volume and protein content, which alludes to the difficulty in separating protein quality (SDS sedimentation) from protein quantity. For the temperature- or humidity-dependent time properties, the correlation coefficients, while significant in most cases, was comparatively small, never exceeding 0.5 in absolute magnitude.

The relationship between the PLS2 loading for Factor 1 and protein content is readily seen in a plot of the sample scores for Factors 1 and 2 (Fig. 3). The size (diameter) of the symbol is linearly proportional to the sample's protein content, with the smallest and largest sizes corresponding to 9.09 and 18.27% protein, respectively. Most evident from this plot is that the Factor 1 scores increase with decrease in protein content. This gradient is not apparent in the scores associated with Factors 2 or higher.
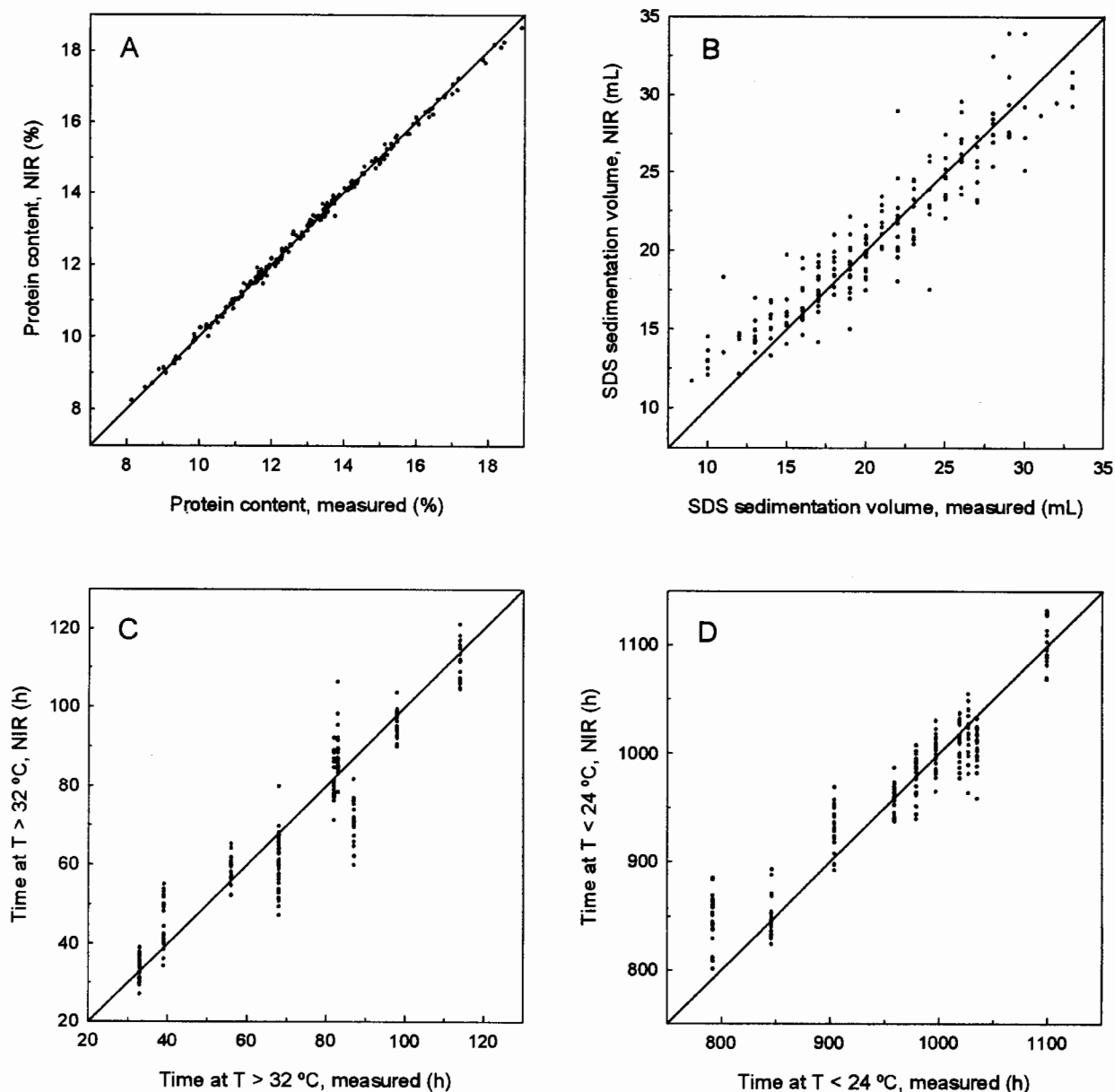

Fig. 4. Test set results for NIR PLS1 modeling of four constituents or conditions. **A,** protein content; **B,** SDS sedimentation volume, **C,** time during grain fill period in which the temperature exceeded 32°C; and **D,** time during grain fill period in which the temperature was lower than 24°C. Model results are summarized in Table III.

Results of PLS1 models are summarized in Table III. Compared with models in which no spectral pretreatment was performed (results not shown), application of a second-derivative tended to improve model performance by reducing the RMSD and the number of factors. The number of factors ranged from 9 (for protein content and SDS sedimentation volume) to 12 (for the humidity-related time conditions). Aside from protein content, RPD values ranged from 2.0 for $t(T > 35°C)$ (poorest model performance) to 3.3 for $t(T > 30°C)$ (best model performance). It appears that NIR reflectance is sensitive to conditions of protein quality (SDS sedimentation volume) and, indirectly, to the environmental conditions that affect it; however, there is no sharp demarcation between a temperature that promotes grain development and one that potentially places the plant in a heat shock condition. Recent research by S. B. Altenbach et al (*unpublished*) on the influence of temperature, drought, or fertilization on the accumulation of genetic transcripts for the gliadins and glutenin subunits in hard red spring wheat (cv. Butte 86) has demonstrated that transcripts have the tendency to appear earlier after anthesis and disappear sooner. However, in contrast to the first hypothesis of Blumenthal et al (1998), Altenbach's team found no evidence to suggest that high temperature (37°C) would result in a change in the proportion of gliadins to glutenins.

The results of the present study suggest the ability of NIR reflectance to monitor quality extends beyond its known ability to measure protein quantity, as demonstrated by partial correlation coefficients ($r_{partial}$, Table III) that are statistically significant ($P < 0.001$). However, this extension is least with SDS sedimentation volume, as seen by the large decline in r value when the effect of protein content is removed (i.e., $r = 0.931$ and 0.555 for before and after removal, respectively). The alternative approach of Wesley et al (1999, 2001) is to carefully select nonoverlapping regions of a spectrum that are attributed to the analytes of interest (e.g., gliadin, glutenin), whereupon the spectra of these pure components are either directly measured (preferred) or mathematically derived through a spectral deconvolution procedure applied to well-characterized mixtures of these components. Spectral reconstruction is then performed on the spectra of the unknown samples to determine the proportions of gliadin and glutenin component spectra, thus minimizing any intercorrelation to protein content. Essentially, this curve-fitting procedure is an alternative to the statistical procedure of the current study in which the contribution of protein content is removed by partial correlation analysis. Although model performance was reduced in comparison with partial least squares modeling, the results of Wesley et al corroborate those of our earlier study (Delwiche et al 1998), in that wheat protein quality, as defined by the levels of gliadin and glutenin, can be monitored by NIR reflectance. In the present study, environmental factors that influence wheat quality are also shown to be indirectly measurable by NIR reflectance. Plots of NIR-modeled vs. measured values of the test set samples for protein content, SDS sedimentation volume, $t(T > 32°C)$, and $t(T < 24°C)$ are shown in Fig. 4 (plots of the other attributes are omitted for the reason of their similarity to the ones shown). In the protein content and SDS sedimentation plots (Fig. 4A and B), values are evenly clustered around the (45°) line of zero model error. The dispersion of the values is greater for the temperature-dependent time parameters (Fig. 4C and D), as shown by the circumstances in which all NIR-predicted values corresponding to samples from a geographical location (forming a vertical cluster of points) were skewed to one side of the 45° line.

Multiple linear regression modeling is summarized in Table IV. In reference to Table IV, absorption bands are associated with wavelengths whose coefficients were negative. Although these models were generally lower in performance than corresponding PLS1 models, their relative simplicity (four terms as opposed to 9–12 factors) may offset this detriment. In contrast to the PLS1 models, model performances for the humidity-dependent time properties were noticeably lower than for the temperature-dependent time properties. Most of the wavelengths selected for the temperature-dependent time property models were from the 1,200–1,900 nm region, a region

predominated by first and second overtone vibrations of CH, OH, and NH (Miller 2001). The second-derivative gap size associated with the wavelength terms for these models typically had a range of 10–16 nm. Some of these wavelengths are readily understandable in terms of known absorbers (e.g., 2,170 nm for protein). Others are not easily interpretable, although they should be useable, given the lack of difference between the standard error of calibration (not shown) and the prediction set's standard error of performance.

CONCLUSIONS

Through partial least squares and multiple linear regression modeling, NIR reflectance on ground wheat has been shown to be sensitive to the environmental conditions that prevailed during the period of grain development. SDS sedimentation volume and the number of hours during the period of grain fill in which temperature was above a specified temperature (30, 32, and 35°C) or below 24°C were reasonably well (RPD range of 2.0–3.3) modeled by PLS1. Likewise, PLS1 models for the number of hours during grain fill at extreme humidity conditions (<40% rh or >80% rh) were also reasonable (RPD = 2.5 and 2.8). With scatter correction and second-derivative spectral pretreatments, the number of PLS1 factors had a range of 9–12, depending on the constituent or property modeled. Multiple linear regression models consisting of four second-derivative terms produced reasonable models for SDS sedimentation volume and two of the temperature-dependent time parameters [$t(T > 30°C)$ and $t(T > 32°C)$]. For both PLS and multiple linear regression models, success in modeling extended beyond an intercorrelation between the modeled property and protein content. With additional research, NIR spectroscopy has the potential for becoming a tool to be used by the wheat industry for assessing the cultural environment under which the purchased grain was produced.

ACKNOWLEDGMENTS

LITERATURE CITED

American Association of Cereal Chemists. 2000. Approved Methods of the AACC, 10th ed. Method 56-70. The Association: St. Paul, MN.

Blumenthal, C., Stone, P. J., Gras, P. W., Bekes, F., Clarke, B., Barlow, E. W. R., Appels, R., and Wrigley, C. W. 1998. Heat-shock protein 70 and dough-quality changes resulting from heat stress during grain filling in wheat. Cereal Chem. 75:43-50.

Blumenthal, C. S., Barlow, E. W. R., and Wrigley, C. W. 1993. Growth environment and wheat quality: The effect of heat stress on dough properties and gluten proteins. J. Cereal Sci. 18:3-21.

Busch, R. H., Shuey, W. C., and Frohberg, R. C. 1969. Response of hard red spring wheat (*Triticum aestivum* L.) to environments in relation to quality characteristics. Crop Sci. 9:813-817.

Ciaffi, M., Tozzi, L., Borghi, B., Corbellini, M., and Lafiandra, D. 1996. Effect of heat shock during grain filling on the gluten protein composition of bread wheat. J. Cereal Sci. 24:91-100.

Daniel, C., and Triboi, E. 2000. Effects of temperature and nitrogen nutrition on the grain composition of winter wheat: Effects on gliadin content and composition. J. Cereal Sci. 32:45-56.

Delwiche, S. R., Graybosch, R. A., and Peterson, C. J. 1998. Predicting protein composition, biochemical properties, and dough-handling properties of hard red winter wheat flour by near-infrared reflectance. Cereal Chem. 75:412-416.

Faridi, H., and Finley, J. W. 1989. Improved wheat for baking. CRC Crit. Rev. Food Sci. Nutr. 28:175-209.

Fisher, L. D., and van Belle, G. 1993. Biostatistics: A Methodology for the Health Sciences. Pages 510-512 in: Wiley-Interscience: New York.

Graybosch, R. A., Peterson, C. J., Baenziger, P. S., and Shelton, D. R. 1995. Environmental modification of hard red winter wheat flour

protein composition. J. Cereal Sci. 22:45-51.

Greenspan, L. 1977. Humidity fixed points of binary saturated aqueous solutions. J. Res. Nat. Bur. Stand. 81A:89-96.

Lookhart, G. L., and Finney, K. F. 1984. Polyacrylamide electrophoresis of wheat gliadins: The effect of environment and germination. Cereal Chem. 61:496-499.

Luo, C., Branlard, G., Griffin, W. B., and McNeil, D. L. 2000. The effect of nitrogen and sulphur fertilisation and their interaction with genotype on wheat glutenins and quality parameters. J. Cereal Sci. 31:185-194.

MacRitchie, F., and Gupta, R. B. 1993. Functionality-composition relationships of wheat flour as a result of variation in sulphur availability. Austr. J. Agric. Res. 44:1767-1774.

Martens, H., and Naes, T. 1989. Multivariate Calibration. Pages 73-236. Wiley: Chichester, England.

Miller, C. E. 2001. Chemical principles of near-infrared technology. Pages 19-37 in: Near-Infrared Technology in the Agricultural and Food Industries, 2nd Ed. P. C. Williams and K. H. Norris, eds. Am. Assoc. Cereal Chem.: St. Paul, MN.

Osborne, B. G., and Fearn, T. 1986. Near Infrared Spectroscopy in Food Analysis. Pages 117-161. Longman Scientific and Technical: Harlow, England.

Paredez-Lopez, O., Corravioubas-Alvarez, M., and Barquin-Carmona, J. 1985. Influence of nitrogen fertilization on the physicochemical and functional properties of bread wheats. Cereal Chem. 62:427-432.

Peterson, C. J., Graybosch, R. A., Baenziger, P. S., and Grombacher, A. W. 1992. Genotype and environment effects on quality characteristics of hard red winter wheat. Crop Sci. 32:98-103.

Randall, P. J., and Moss, H. J. 1990. Some effects of temperature regime during grain filling on wheat quality. Austr. J. Agric. Res. 41:603-617.

Stone, P. J., and Nicolas, M. E. 1995. A survey of the effects of high temperature during grain filling on yield and quality of 75 cultivars. Aust. J. Agric. Res. 46:475-492.

Stone, P. J., Gras, P. W., and Nicolas, M. E. 1997. The influence of recovery temperature on the effects of a brief heat shock on wheat. III. Grain protein composition and dough properties. J. Cereal Sci. 25:129-141.

Wesley, I. J., Larroque, O., Osborne, B. G., Azudin, N., Allen, H., and Skerritt, J. H. 2001. Measurement of gliadin and glutenin content of flour by NIR spectroscopy. J. Cereal Sci. 34:125-133.

Wesley, I. J., Uthayakumaran, S., Anderssen, R. S., Cornish, G. B., Bekes, F., Osborne, B. G., and Skerritt, J. H. 1999. A curve-fitting approach to the near infrared reflectance measurement of wheat flour proteins which influence dough quality. J. Near Infrared Spectrosc. 7:229-240.

Williams, P. C., and Sobering, D. C. 1993. Comparison of commercial near infrared transmittance and reflectance instruments for analysis of whole grains and seeds. J. Near Infrared Spectrosc. 1:25-32.